© EARSeL and Warsaw University, Warsaw 2005. *Proceedings of 4th EARSeL Workshop on Imaging Spectroscopy. New quality in environmental studies*. Zagajewski B., Sobczak M., Wrzesień M., (eds)

DETECTION OF BIOTIC STRESS (*VENTURIA INAEQUALIS*) IN APPLE TREES USING HYPERSPECTRAL ANALYSIS

Stephanie Delalieux, Jan van Aardt, Wannes Keulemans, and Pol Coppin

Katholieke Universiteit Leuven, Department of Biosystems: Measure, Model, and Manage Bioresponses (M3-BIORES). Vital Decosterstraat 102, 3000 Leuven, Belgium; Stephanie.Delalieux@agr.kuleuven.ac.be

ABSTRACT

The potential yield of (capital)-intensive multi-annual crops (e.g., fruit) is seldom harvested in reality. A targeted monitoring and modelling of the growth processes in such agricultural production systems could enable an early detection and treatment of production limiting factors, thereby optimising yield. In Belgium, as in all temperate regions, scab stress caused by the ascomycete *Venturia Inaequalis* causes the most important stress in apple orchards. The objectives of this study were (i) to determine if *Venturia inaequalis* leaf infections could be differentiated from healthy leaves in both resistant and susceptible cultivars using hyperspectral spectroradiometer data, (ii) to gauge at which developmental stage *Venturia inaequalis* infections could be detected and (iii) to identify wavelengths or spectral regions that best differentiate between infected and healthy leaf material. The first objective was related directly to the scientific research question of whether or not infected leaves resulted in a spectral response different from healthy leaves. The second objective addressed the question of whether or not hyperspectral data could serve as part of an early warning system to biotic stress in apple orchards, while the last objective defines the practical implication of such a spectral warning system. Partial Least Squares Discriminant Analysis (PLSDA) was used as classification technique. This technique compresses the dimension of the hyperspectral reflectance dataset, followed by a discriminant classification. Results suggested that good predictability could be achieved when classifying infected plants based on hyperspectral data using PLSDA. Furthermore, a band reduction technique based on logistic regression was used to select the hyperspectral bands that best define differences among treatments. This study showed that the spectral domains centered around 1500nm and the visible region (well-developed infection stage) were best suited to differentiate between infected and healthy plants.

INTRODUCTION

The structure and physiological status of a plant is represented by reflectance patterns (i). Incident energy of the sun is partly reflected, transmitted and absorbed by the plant. The amount of reflected light depends on an amount of leaf-related factors, such as external morphology, internal structure, concentration, and internal distribution of biochemical components. Each possibility to monitor these factors in an ongoing or periodic manner by means of teledetection offers the potential to model plant production processes, and therefore also to steer processes by means of adapted management measures. Studies have shown that non-intrusive techniques are essential for capturing data in the continuous manner necessary for monitoring vegetative production systems.

Much remote sensing research has been done at plant leaf-level to ascertain the amount of stress in plants. Most studies focused on physiological changes and how these changes alter the interaction of light with the foliar medium (ii, iii). The most common and widespread change occurs in the proportion of light-absorbing pigments, most notably chlorophylls *a* and *b* which absorb light in the 430–660nm region (iv). Investigators have observed differences in reflectance due to stress-induced changes in pigment concentration in the *green peak (525-605nm)* and along the *red edge (~750nm)* (v, vi, vii, viii).

Apple scab stress manifests at different stages. As can be seen in Figure 1, the primary infection is initiated by the release of ascospores. These spores penetrate the leaf cuticle, causing the formation of a flat mycelium between the cuticle and the epidermal cell walls. Sporulation occurs by means of conidiophores, which penetrate the cuticular layer of the leaf surface and release conidia (ix). Until sporulation, *V. inaequalis* grows without killing host cells. First, olive green spots will appear on the leaf surface, followed by typical necrotic scab lesions. At the end of the vegetation period, the fungus switches to saprophytic growth in the dead leaf tissue and forms fruit bodies (pseudothecia), which overwinter and give rise to ascospores the following spring. (x)

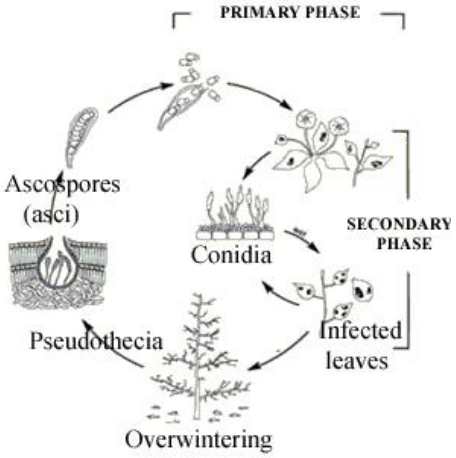

Figure 1 : Apple scab disease cycle

Little is known about the host-pathogen interaction of apple scab. Resistant apple cultivars generally activate their defense mechanism after infection, which initiates a sequence of metabolic reactions. In some cases it will cause a hypersensitive reaction of the plant cell that has been penetrated by the pathogen, resulting in cell death. When this happens, the pathogen will in many cases die, due to toxic substances accumulating in the dead plant cell. Lateur (xi) listed the most common metabolic modifications a) the reinforcement of the mechanical barrier consisting of a pecto-cellulose wall by incorporation of polysaccharides (callose), phenolic polymers (lignine), polyesters (suberine) and proteins, b) the stimulation of defense enzymes, and c) the production of defense and pathogenesis related proteins, such as chitinases, glucanases, and thaumatine-like proteins. All these factors will influence the reflectance spectra of the vegetation. Figure 2 shows the influence of selected vegetation compounds on a reflectance spectrum from plant material (xii, xiii, xiv, xv, xvi).

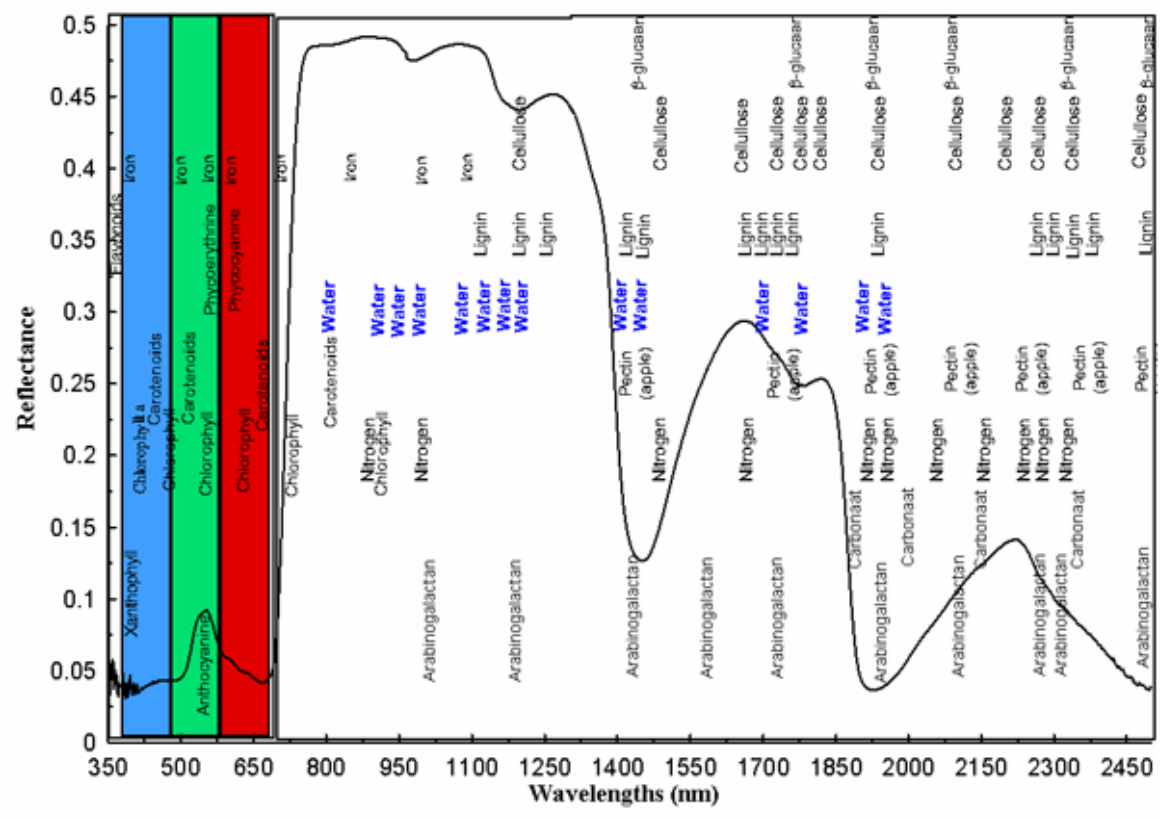

Figure 2: Vegetation spectrum and reflectance (4)

Hyperspectral data were used in this study because of their contiguous nature and the abundance of narrow wavelength bands in the electromagnetic reflectance spectrum. These factors render hyperspectral data as a result of which they are very sensitive to subtle variations in energetic reflectance. This unique sensitivity is in turn needed to detect anomalies in the normal growth processes of plants, but comes at the cost of very large datasets, so that processing the data becomes a real challenge regarding computational power, as well as information redundancy removal, relevant information extraction and modeling accuracy. Statistical techniques therefore have to be especially robust, given such large multivariate datasets. Most of the traditional methods of classification and band reduction rely on assumptions of normality. Unfortunately, the data in this research did not fulfill this assumption. This necessitated the use of specialized techniques to determine if a difference between exposed and non-exposed leaves existed. “Partial least squares” was applied to reduce the data dimensionality (xvii). Finally, discriminant analysis was used on those results to classify the data. A technique based on logistic regression per wavelength was used to maintain statistical robustness, while defining wavelengths that resulted in the best separability between healthy and infected plants.

METHODS

Leaf-level reflectance data were obtained for two different apple cultivars in this study. The plant material was cloned *in vitro*, rooted, and acclimatized before being used in experiments. A cultivar susceptible to scab infection, Braeburn was selected as well as a resistant cultivar, Rewena. The objective was to recognize the different reaction mechanisms in plants and recognize their developmental stage.

Vegetative material

Each of both cultivars (Braeburn and Rewena) contained thirty plants: ten infected, ten placebo and ten untreated, or control plants. A standard infection method was used to infect the plants. The plants were inoculated with a suspension of 150,000 conidiospores per milliliter. The suspension was obtained by rinsing spores of leaves exhibiting visible scab symptoms. The concentration was

then counted in a Neubauer counting chamber to control and adjust the solution. Actively growing plants with 3 - 4 unfolded leaves were infected by spraying them with the suspension and placing them in 100% relative humidity for 48 hours. Afterwards, the plants were placed in an environment with a relative humidity of approximately 60%. Placebo plants were treated the same way, but in this case the suspension contained no bacteria.

Data collection

All spectral signatures were collected by spectral scanning between 350 and 2500 nm with an Analytical Spectral Devices Inc. (ASDI) FieldSpec Pro JR spectroradiometer (xviii). Leaf reflectance spectra were captured by using a plant probe, which measures reflectance spots of 10 mm diameter and comes with a 1500 hour stable halogen bulb. The sampling interval over the 350–1050 nm range is 1.4 nm with a spectral resolution of 3 nm (band width at half maximum). The sampling interval and the spectral resolution are approximately 2 nm and 30 nm respectively for the 1050–2500 nm range. The results are then interpolated by the ASD software to produce readings at each nanometer interval.

The measurements were initiated ten days after infection and were repeated every four days until scab symptoms were clearly visible with the eye, and leaves became to fall off. The two youngest vulnerable leaves from each plant were masked during infection. Two samples were taken during every measurement from each of these leaves, so that a good representation of the whole leaf was obtained. Photographs were taken so that later analyses could be linked with the degree of infection. The leaves were classified according to the scale of *Chevalier et al.* (xix). It contains six classes of symptoms: class 0 = no symptoms; class 1 = hypersensitivity (pinpoint pits); class 2 = resistance, chlorosis, and necrosis, either together or separate, without sporulation; class 3a = weak resistance, chlorosis, and necrosis, either together or separate, with slight sporulation; class 3b = weak susceptibility, chlorosis, and necrosis, either together or separate, with sporulation; and class 4 = susceptibility, abundant sporulation without chlorosis and necrosis, either together or separate.

Data analysis

The gathered leaf-level reflectance data were subjected to partial least squares discriminant analysis and logistic regression on a wavelength basis. A short description of both techniques is given below.

- Partial least squares discriminant analysis (xx)

  Linear discriminant analysis would traditionally have been the most appropriate technique to classify the data, given that the data were normally distributed. The dataset was tested for normality using the Shapiro-Wilk normality test and by calculating data skewness, both of which indicated a non-normal data distribution for many wavelength variables. A formal linear discriminant analysis also cannot be performed due to the large number of variables in the available training dataset relative to the amount of measurements taken. A reduction in data dimensionality therefore was needed to avoid overfitting before continuing with the classification. PLSDA therefore was preferred to analyze the data. The first step in this technique is a dimension reduction by using partial least squares (PLS). PLS is comparable to the commonly used dimension reduction technique of principal component analysis (PCA), with the important difference being that PLS explains both sample variation and response variation. In contrast with PCA, PLS components are chosen such that the sample covariance between the response and a linear combination of the predictors is maximized. This criterion for PLS is more sensible since there is not an *a priori* reason why constructed components with large predictor variation should be strongly related to the response variable (xxi). A component with a small predictor variance could be a better predictor of the response classes, a fact which is not taken into account in PCA.

  The main objective of partial least squares is to build a model which relates the response variables to the factor scores multiplied by their loadings. The factor scores, in turn, are lin-

ear combinations of the original predictor variables, resulting in no correlation between the factor score variables used in the predictive regression model (xxii).

$$Y=TQ+E \quad (1)$$

where Y = n x m response variables matrix

T = factor score matrix (=predictor variables x weights)

Q = coefficients matrix (=loadings for T)

E = n x m noise term

The second step in the PLSDA technique involves a classification using linear discriminant analysis (LDA). LDA is well-known as a classification technique based on the gross variability 'within groups' and 'among groups'. The combination of PLS and LDA therefore results in a dimension reduction as well as a classification outcome.

A cross-validation 'random subset' method was chosen in order to evaluate the model obtained by PLSDA. The basic precept behind this model validation technique is that a data subset (test data) is removed before training begins. The performance of the selected model then can be tested on the new test data. Here, the dataset was divided in 5 parts and was iterated 20 times.

The number of latent variables where then specified that had to be retained in the model, which relates to the percent variance captured by the model in X and Y. Score plots, afterwards were made for both cultivars at each different measurement period. These score plots represented the predicted outcomes obtained by PLSDA.

Overall sensitivity and specificity of the model can be obtained by calculating the C-index. This index represents the area under the Receiver-Operator Characteristic (ROC) curve and gives an estimate of the discriminatory performance of the system. The vertical axis by convention is sensitivity, also known as true positive rate (or the ability to predict an event correctly). The horizontal axis is (1-specificity) otherwise known as the false positive rate (or the proportion of predicted event responses that were observed as nonevents). The curve gives the true positive rate that can be obtained at the cost of each possible false positive rate (xxiii). A numerical measure of the accuracy of the model can be obtained from the area under the curve, where an area of 1.0 signifies near perfect accuracy, while an area of less than 0.5 indicates that the model is worse than just random. C-values above 0.8 are generally accepted to represent significant discriminative models (xxiv).

The PLS analyses and DA classifications were performed in Matlab 6.5 using the PLS toolbox written by Barry M. Wise et al. (xxv).

- Logistic Regression per Wavelength

The band selection method used in this research is based on logistic regression, which is an alternative procedure to discriminant analysis. In contrast to discriminant analysis, logistic regression analysis does not make any assumptions regarding the distributions of the independent variables. Unfortunately, it was not possible to perform logistic regression on the dataset as a whole, because of the large independent-dependent variable ratio. Per-wavelength logistic regression therefore was the method of choice in this study.

The approach presented here was able to detect significant differences in reflectance between stressed and non-stressed leaves. The categorical response variables had only two values. Leaves were designated as either stressed (0) or not stressed (1). Logistic regression was performed for each wavelength for this reason, as well as to avoid the normality assumption. Summary statistics are of direct interest, and often are reported as percentages or proportions. Logistic regression models these proportions as function of explanatory variables and provides a solution to the principle questions of which variables (wavelengths) are important in differentiating between stressed and healthy leaves.

The logistic regression model is a non-linear transformation of the linear regression. It is a statistical technique for making predictions when the dependent variable is a dichotomy, and the independent variables are continuous and/or discrete. A simple logistic regression model can be stated as follows (xxvi) :

$$Y_i = E\{Y_i\} + \varepsilon_i \, Y_i \qquad (2)$$

where Y_i are independent Bernouilli random variables with expected values $E\{Y_i\}$, defined by:

$$E\{Y_i\} = \frac{\exp(\beta_0 + \beta_1 X_i)}{1 + \exp(\beta_0 + \beta_1 X_i)} \qquad (3)$$

The X observations are known constants.

The C-index was extracted from the statistical analysis. This index was also used in the PLSDA technique.

The logistic regression (xxvii) based analyses are programmed in the computational language 'R', available as free software (xxviii).

RESULTS AND DISCUSSION

The normality of the dataset first was tested to verify which statistical analysis best should be applied to the hyperspectral dataset. Statistical techniques based on normality assumptions could not be used in this study. This conclusion was based on results from the Shapiro-Wilk test, as well as the data skewness exhibited by Rewena cultivar data from the sixth weekly measurement (Figure 3). High p-values (>0.2) from the Shapiro–Wilk analysis are indicative of a normal distribution, as well as relatively low skewness values (< 0.5). Unfortunately, 1069 wavelength variables were not normally distributed in this specific case (Figure 3), based on the defining parameters mentioned above.

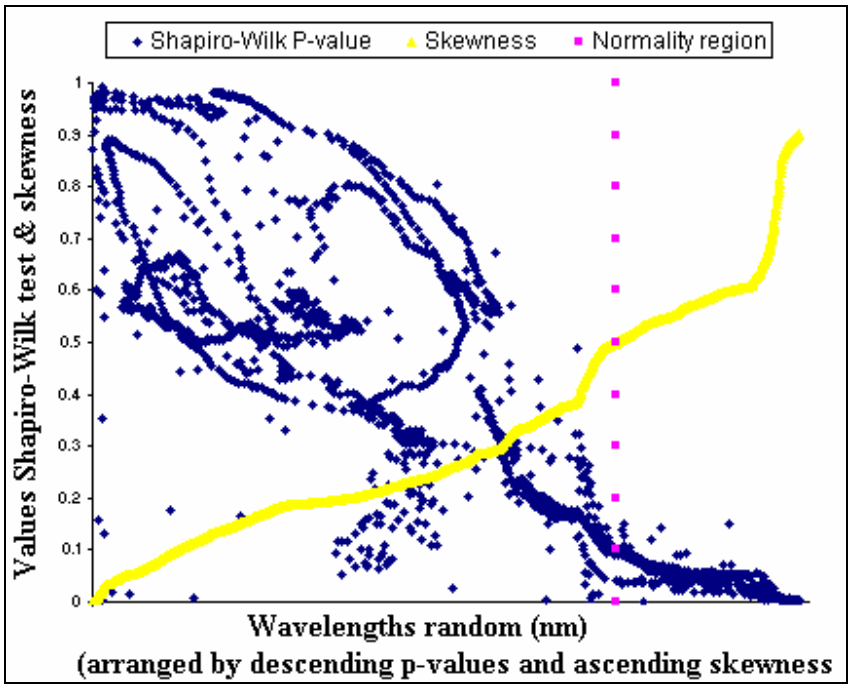

Figure 3: Results of the Shapiro-Wilk test for normality and skewness-values. It is clear that a large number of wavelength variables were not normally distributed, with associated skewed data.

PLSDA classification

PLS was used, as part of the PLSDA technique, to develop a model that predicts the class number (0 or 1) for each sample. The model will not perfectly fit 0 or 1, but a limit can be set above which the sample is estimated as one and below which it is designated as zero. Score plots represented the predicted outcomes obtained by PLSDA.

The following figures illustrate score plots at the beginning and end phase of the experiment for the susceptible Braeburn cultivar as well as for the resistant Rewena cultivar. Only the predictions for Y that were classified as zero are shown here.

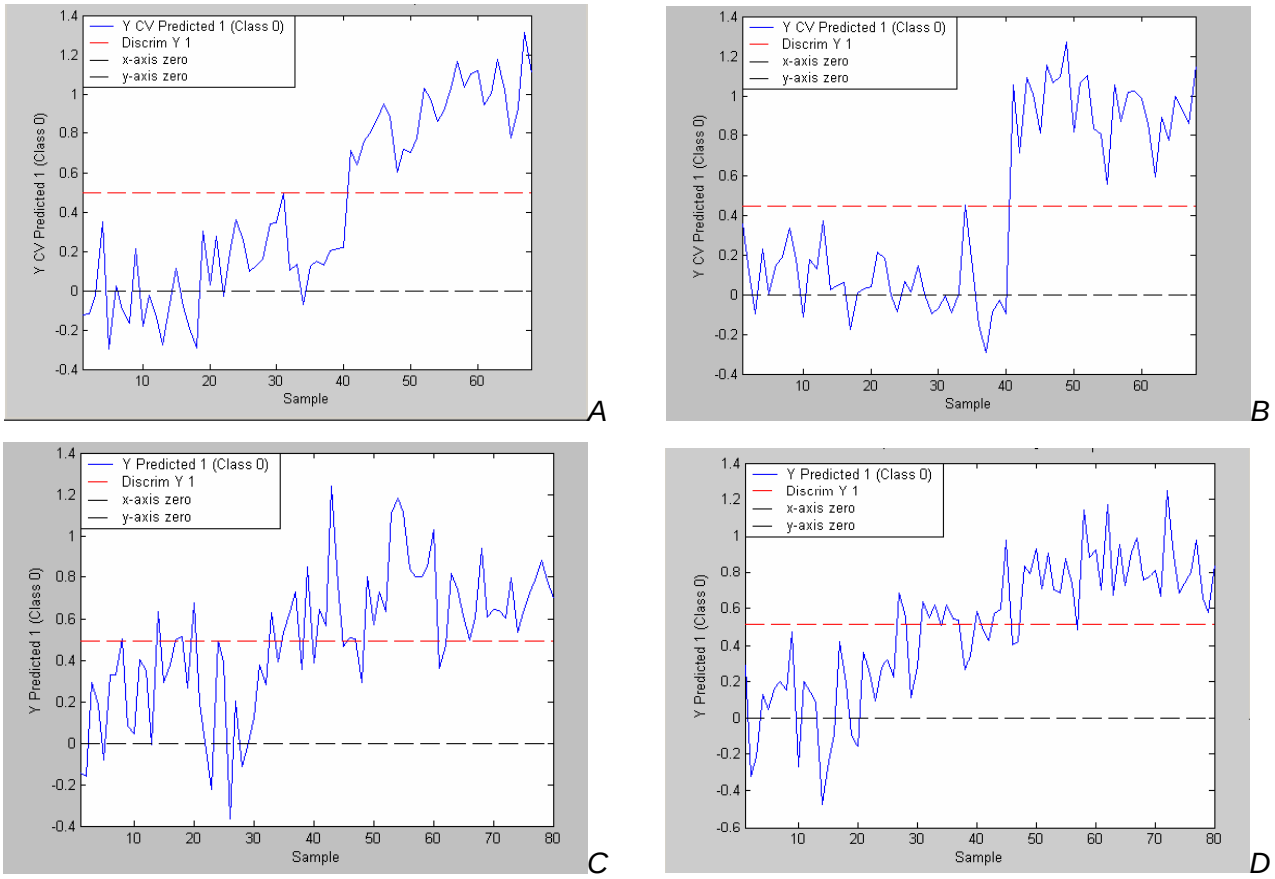

Figure 4: Score plots (Y values predicted as '0 class') obtained by PLSDA for the susceptible Braeburn (A&B) and resistant Rewena (C&D) cultivar at the start (week1; A&C) and at the end (week6; B&D) of the experiment.

Notice that the predicted scores for the susceptible Braeburn cultivar (Fig.4 A&B) performed better than those for the resistant Rewena cultivar (Fig. 4 C&D). As expected, it can be concluded that a bigger difference existed between the healthy and infected leaves of the susceptible cultivar as opposed to the resistant cultivar.

Loading plots were made for each selected latent variable to determine which wavelength or wavelength domain had the largest loading in the analysis. Figure 5 shows a loading plot of latent variable 1 from the Braeburn cultivar dataset for the first as well as last measurement. It is clear that the infrared domain (750-2500 nm) received the largest variable loadings immediately after infection, while the visible domain from 650-700 nm played an important role in the ability to separate spectra from healthy and infected leaves 31 days after infection. Other latent variables (LV 2, 3, 4, 5 and 6) from day 10 focused on the domain around 550 nm and 750 nm as most important wavelength regions. Accordingly, the wavelengths around 550 nm, 650 nm and 750nm were well represented based on the impact loadings on other latent variables (LV 2-7) from day 31. It can be concluded from these results that the infection had a significant impact on the moisture level of the susceptible leaves during the first days after exposure. This deduction is based on the factor load-

ings and implied presence of near-infrared regions in the analysis. However, when the scab infection was more pronounced (e.g., at 31 days), the importance of visible regions increased based on variable loadings. This was due to structural cell changes in the leaves induced by the biotic stress, while visible chlorotic lesions appeared. Since these lesions were evident to the naked eye, visible regions became increasingly important in the statistical differentiation. An example of a control and infected leaf of the susceptible Braeburn cultivar are shown in Figure 6.

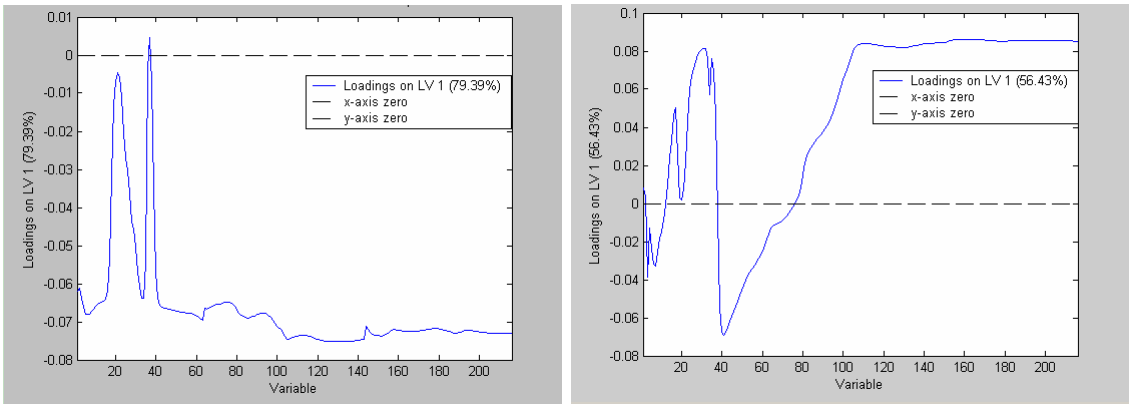

Figure 5: Loading plots for Braeburn 10 days after infection (left) and 31 days after infection (right).

Figure 6: A picture of an untreated control (left) and infected leaf (right) of the susceptible Braeburn cultivar, taken 31 days after infection.

ROC curves of the 'Braeburn' and 'Rewena' cultivars, measured 31 days after infection, are shown in Figure 7. It is clear that the discriminatory performance of the PLSDA method is markedly better for the susceptible cultivar, as opposed to the resistant Rewena cultivar.

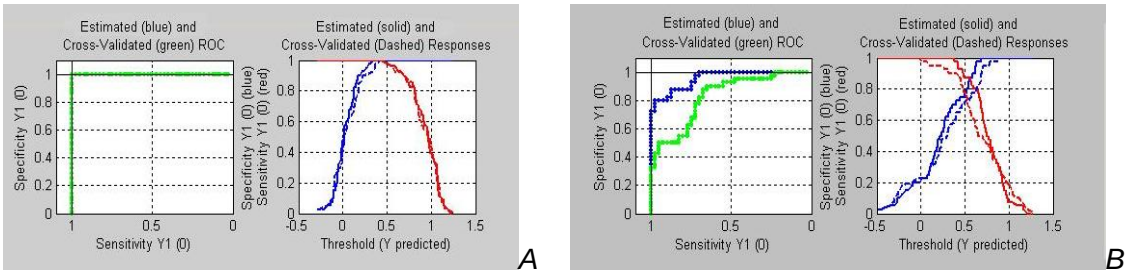

Figure 7: ROC curves of the susceptible cultivar (fig A) and the resistant cultivar (fig B); measurements taken 31 days after infection.

Logistic regression band reduction

The spectral bands that best differentiated between classes (infected-healthy) were selected by the logistic regression technique as discussed earlier. The "logistic" distribution is an S-shaped function. The distribution constrains the estimated predicted values to lie between 0 and 1 and can be interpreted as probabilities of stress.

The graph in Figure 8 shows the relationship between the reflectance values of the stressed and healthy leaves and the predicted probability of stress at one specific wavelength (707nm).

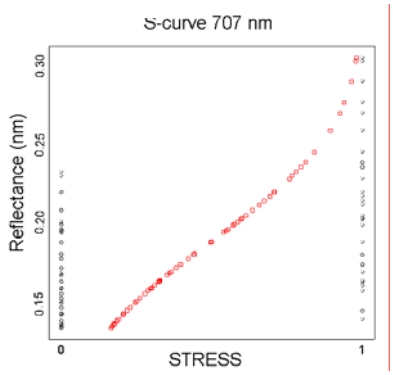

Figure 8: S-shaped probability curve based on logistic regression at wavelength 707nm.

The C-index, an estimation of the discriminatory performance of the system can be calculated with the function 'lrm' in the computational language 'R'. Figure 9 shows the value of the C-index along with other indicative statistical parameters (R^2 and 1-P). It is clear from Figure 9 that R^2 values were not useful to validate logistic models. It was furthermore concluded that 1-P values did not represent adequate significance indicators for separability of spectra of healthy and infected leaves. Figure 9A represents the results of the per-wavelength logistic regression from the resistant Rewena cultivar measured 10 days after infection, while 9B represents the results from the same cultivar 31 days after infection. It was concluded that no significant spectral differences existed between healthy and stressed leaves for the resistant Rewena cultivar at the start and end of the measurement period. This was based on C-index values of 0.5 or less, which is indicative of no discrimination. Figures (C&D) represent the results of the logistic based reduction technique for the susceptible cultivar 'Braeburn'.

The statistical analysis of the susceptible 'Braeburn' cultivar datasets (Figure 9 C&D) clearly indicated regions of the spectrum where the difference in reflectance between stressed and unstressed leaves was significant. The infection had a distinct influence on the reflectance spectra in the region around 1500nm, which is well-known for its water absorption features and cell wall structure. Infection by *V. inaequalis* has a profound effect on the distribution pattern of water and dissolved solute within leaf veins. Water loss from lesions furthermore is a factor in a solute transport system that has been proposed as a method by which *V. inaequalis* acquires nutrients for growth under the cuticle (xxix). Measurements taken 31 days after infection (Figure 9D) showed a high C-index value in the visible region and the red edge. The first spectral region of interest (556-660nm) corresponds to the regions of chlorophyll absorption and also can be seen with the visible eye due to chlorotic and necrotic lesions. The reduction in chlorophyll and carotenoid content was regarded as a direct response to an irreversible disorganization of the chloroplasts. Scab spores growing into stomatal cavities can injure cells or block openings, thus preventing the exchange of CO_2, O_2 and water vapor. Scab infection hereby indirectly affects the photosynthesis of the infected leaf, which results in a change in chlorophyll content, the most important component in photosynthesis. The red-edge region (685-715 nm) has been shown to be another stress sensitive spectral range. Reflectance differences are also found to be significantly different in the region between 760 and 800 nm. These regions conform to expectations based on biophysical knowledge of the effect of biotic stress upon leaves (28).

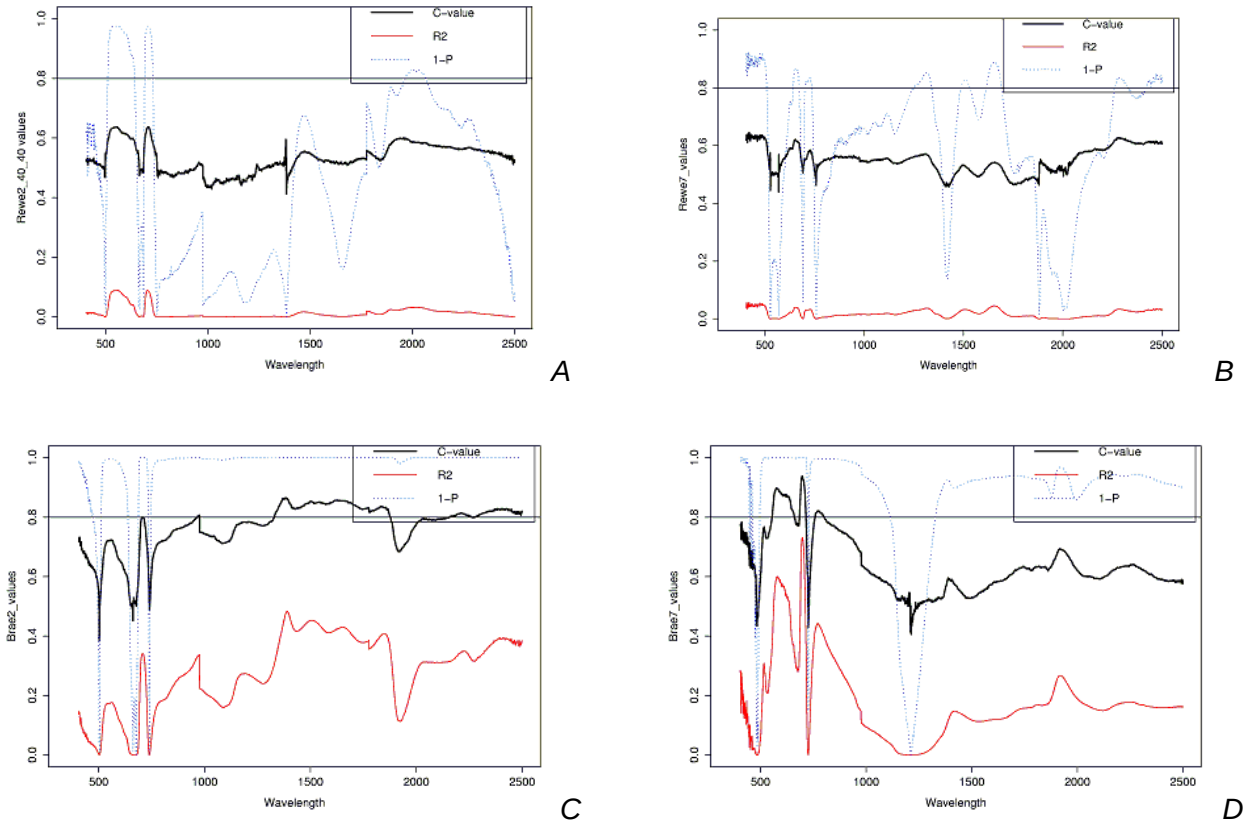

Figure 9: Statistical output of the logistic based technique for the resistant Rewena cultivar (A&B) and the susceptible Braeburn cultivar (C&D) measured 10 days after infection (A&C) and 31 days after infection (B&D). C-index values, R^2, and (1-P) values are represented.

The Rewena cultivar is known to be resistant to scab. In Figure 9, only the results of the first and the last measurements were shown. However, in the analyses of all measured time periods, it was remarkable that 14 days after infection the C-index reached a value of approximately 0.75 for the 1570 nm region. This value decreased after 21 days, resulting in no significant differences between healthy and infected leaf spectra 31 days after infection. The high C-values were attributed to the activation of the resistance mechanism. The causality of these effects needs to be investigated by extended further research.

CONCLUSIONS

Results suggested that acceptable predictability could be achieved using hyperspectral data and partial least squares discriminant analysis to classify susceptible apple plants infected with *Venturia inaequalis.* The two statistical techniques (PLSDA and logistic regression) resulted in the selection of similar spectral domains (1300 -1700nm and 550-750nm in well-developed stage) for the differentiation between spectra from healthy and infected leaves.

Inoculation of apple leaves with a scab suspension resulted in leaf chlorosis in the case of the susceptible Braeburn cultivar. Stress therefore could be detected effectively at the end of the measurement cycle using visible wavelength bands. The statistical methods also were capable of detecting differences in reflectance responses before they were visible to the naked eye.

In conclusion, the spectral domain between 1300-1700 nm proved to be the most effective region to separate stressed from healthy leaves immediately after infection. As the scab infection progressed, the visible wavelengths became more important in the case of the susceptible Braeburn cultivar. The resistant Rewena cultivar appeared to have a short time-span 14 days until 21 days after infection where a difference could be detected between healthy and infected leaves. Although

this sudden appearance/disappearance of stress detection ability was attributed to the activation of the resistance mechanism, further research is needed to discover the causality of these effects.

These results suggest that it could be possible to detect biotic plant stress using remote sensing. This is of significant importance to the agricultural market. For instance, an early warning system, based on spectral inputs, would be an ideal solution to the enforced reduction of pesticide use. Farmers only have to apply pesticides when abnormalities are detected in the normal growth pattern of the plants. This study therefore has important implications for biotic stress detection and management in apple orchards. At-satellite measurements will enable managers to obtain frequent hyperspectral coverage of large areas, thereby making continuous monitoring of biotic stress in apple orchards a reality.

FUTURE RESEARCH

Further research on this topic will concentrate on alternative statistical methods, such as tree-based modeling and neural networks. In-depth research also is required in order to define the exact biophysical mechanisms that are linked to specific wavelength ranges.

ACKNOWLEDGEMENT

This project was funded by the Katholieke Universiteit Leuven, Department of Biosystems. The authors further would like to acknowledge funding for preliminary research (Hypercrunch project) provided by the Belgian Science Policy Office (BELSPO; Carine Petit and Joost Vandenabeele).

References

i Penuelas J & Filella I, 1998. Visible and near-infrared reflectance techniques for diagnosing plant physiological status, Trends in Plant Science 3: 151-156

ii Gausman H, 1985. Plant leaf optical parameters in visible and near-infrared light. Graduate studies. Texas Tech Univ. No. 29. (Texas Tech Press, Lubbock, Tex.) 78pp

iii Merzlyak M, Gitelson A, Chivkunova A & Pogosyan S, 2003. Application of reflectance spectroscopy for analysis of higher plant pigments, Russian journal of plant physiology 50: 704-710

iv Zarco-Tejada P, Miller J, Morales A, Berjon A, & Agüera J, 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open canopy tree crops, Remote sensing of environment 90: 463-476

v Smith K, Steven M & Colls J, 2004. Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks, Remote sensing of environment, In Press.

vi Vogelmann T, 1993. Plant tissue optics, Annual Revision of Plant Physiology & Molecular Biology 44: 231-251

vii Carter G.A., 1991. Primary and secondary effects of water content on the spectral reflectance of leaves, American Journal of Botany 78: 916-924

viii Gitelson A & Merzlyak M, 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. Journal of Plant Physiology 148: 495-500

ix Win J, Greenwood D & Plummer K, 2003. Characterisation of a protein from *Venturia inaequalis* that induces necrosis in *Malus* carrying the V*m* resistance gene, Physiological and Molecular Plant Pathology 62: 193-202

x Hahn M, 2003. [online] Pathogenesis-associated differentiation and gene expression in *Venturia inaequalis*, the apple scab pathogen. Available on http://www.uni-kl.de/FB-Biologie/AG-Hahn/Research/venturia%20engl.html [10/03/2005]

xi Lateur M, 2002. Perspectives de lutte contre les maladies des arbres fruitiers à pépins au moyen de substances naturelles inductrices d'une résistance systémique, Biotechnol.Agron.Soc.Environment 6:67-77

xii Curran P, Dungan JL & Peterson D, 2001. Estimatig the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies, Remote sensing of environment 76: 349-359

xiii Gitelson A, Chivkunova A & Pogosyan S, 2003. Application of reflectance spectroscopy for analysis of higher plant pigments. Russian journal of plant physiology 50:704-710.

xiv Merzlyak M, Solovchenko A & Gitelson A, 2002. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit, Postharvest Biology and Technology 27:197-211

xv Penuelas J, Gamon J, Fredeen A, Merino J & Field C, 1993. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves, Remote sensing of environment 48: 135-146

xvi Sims D & Gamon J, 2002. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features, Remote sensing of environment 84: 526-537

xvii Wilson D, Ustin S & Rocke D, 2004. Classification of contamination in salt marsh plants using hyperspectral reflectance. IEEE transactions on geoscience and remote sensing, 42, 1088-1095

xviii ASD, 1999. Analytical Spectral Devices, Inc. Technical guide, 3rd edition, Boulder, Colorado

xix Chevalier M, Lespinasse Y & Renaudin S, 1991. A microscopic study of the different classes of symptoms coded by the Vf gene in apple for resistance to scab (*Venturia inaequalis*). Plant Pathology 40: 249-256

xx Barker M & Rayens W, 2003. Partial least squares for discrimination. Journal of chemometrics, 17: 166-173

xxi Nguyen D & Rocke D, 2002. Tumor classification by partial least squares using microarray gene expression data. Bioinformatics, 18: 39-50

xxii Statistica, StatSoft, 2003. [online] Partial Least Squares (PLS). Available on http://www.bath.ac.uk/mech-eng//auto/textbook/stpls.html [05/04/2005]

xxiii Newson R, 2001. Parameters behind "non-parametric" statistics: Kendall's τ_a, Somers' *D* and median differences. The Stata Journal, 1:1-20

xxiv Boyd S, 2005 [online] ROC Curves Documentation. Available on: http://pops.csse.monash.edu.au/roccurves_doc.html [0.5/04/2005]

xxv Wise B, Shaver J, Gallagher N, Windig W & Koch, 2004. RPLS_Toolbox 3.5 for use with MATLAB™ (Eigenvector Research) 254pp.

xxvi Neter J, Kutner M, Nachtsheim C & Wasserman W, 1996. Applied linear statistical models chapter 14: Logistic regression, poisson regression, and generalized linear models. (McGraw-Hill Companies) 1408pp.

xxvii Harrell F, 2001. Regression modeling strategies. (Springer Verlag, New York) 568pp.

xxviii Center for Computational Intelligence, 2005. [online] The R Project for Statistical Computing. Available on: http://www.r-project.org/

xxix MacHardy W, 1996. Apple Scab: Biology, Epidemiology, and Management. 545pp.